

VU Research Portal

Semantic Support for Quantitative Research

Rijgersberg, H.

2013

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Rijgersberg, H. (2013). *Semantic Support for Quantitative Research*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam].

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Contents

1	Introduction	1
1.1	Quantitative and qualitative research	2
1.2	Limitations of computers in science	3
1.3	Food science research example	4
1.4	The Semantic Web	4
1.5	Virtual Lab e-Science project and Commit project	6
1.6	Research Question	6
1.7	Approach	7
1.8	Contributions	8
1.9	Outline of the thesis	9
1.10	Publications	10
1.11	Cover illustration	10
2	Ontology of Quantitative Research (OQR)	11
2.1	Introduction	11
2.2	Quantitative Research Considerations	12
2.3	Computer support of quantitative research	14
2.4	A quantitative research model	15
2.5	Building a quantitative vocabulary	20
2.6	OQR food science example	21
2.7	The QeSI prototype tool application	24
2.8	Conclusion	26
3	Ontology of units of Measure and related concepts (OM)	35
3.1	Introduction	35
3.2	Related work	37
3.3	Drafting a unified semi-formal description of the domain of units of measure	38
3.4	Description of the domain	39
3.5	Analyzing existing vocabularies of units of measure	43
3.6	Use cases	44
3.7	Design and usage of OM	46
3.7.1	Design of the ontology	47
3.7.2	Modeling issues	50
3.7.3	Functional support provided by ontologies of units	56
3.8	Comparing OM with QUDT	57
3.9	Applying OM	61
3.9.1	Web services and a demo web application	61
3.9.2	The Semantic Calculator	62
3.9.3	Evaluation	66

3.9.4	The OM Excel add-in Rosanne	67
3.10	Discussion and conclusion	70
4	Ontology of computations	75
4.1	Introduction	75
4.2	Illustration of problems	78
4.3	Related work	81
4.4	Requirements	82
4.5	Modeling tabular data in experimental science	84
4.5.1	Classical table representation	85
4.5.2	The semantic table	88
4.5.3	Classical table extended	92
4.6	Modeling computational methods	94
4.6.1	Outline of the approach	94
4.6.2	Illustration of the use of OQR	95
4.6.3	Illustration of the use of OQR in the food science example	97
4.7	Automatically calling external computational methods: stripping and enriching data	100
4.7.1	Stripping and enriching variables in the “mean” example	100
4.7.2	Stripping and enriching variables in the PCA example	104
4.8	Evaluation	108
4.9	Discussion and conclusion	115
5	Annotating quantitative legacy data	119
5.1	Introduction	119
5.2	Problem description	121
5.3	Related work	122
5.4	Materials	124
5.4.1	Datasets	124
5.4.2	Ontology	125
5.5	Approach	126
5.5.1	Tokenization	126
5.5.2	Basic matching: full names and symbols	126
5.5.3	Matching: compounds in OM	126
5.5.4	Matching: compounds not in OM	127
5.5.5	Disambiguation	128
5.5.6	Implementation	130
5.6	Evaluation and analysis	130
5.6.1	Evaluation type and data selection	130
5.6.2	Gold standard creation	131
5.6.3	Results	131
5.6.4	Quantitative analysis	132
5.7	Discussion	133
6	Conclusions	135
6.1	What have we achieved?	135
6.2	The research questions revisited	137
6.2.1	Subquestion 1. What constitutes a quantitative research vocabulary?	137

6.2.2	Subquestion 2. Which tools can be developed to support quantitative research processes?	139
6.2.3	Subquestion 3. How can legacy data be automatically semantically upgraded?	139
6.2.4	Main research question: “How can we support quantitative research processes using formal vocabularies?”	140
6.3	Future outlook.....	140
	Bibliography	151
	Summary.....	159
	Samenvatting	163